

Listing of the Claims:

The following is a complete listing of all the claims in the application, with an indication of the status of each:

Claims 1-22 (Canceled).

- 1 23. (Currently Amended) A method of outlier detection comprising:
 - 2 generating a plurality of synthesized data, each representing a randomly
3 generated state within a given vector space, said generating including a random
4 number generation;
 - 5 receiving a plurality of real sample data, each representing a detected real
6 event as represented in said given vector space;
 - 7 forming a candidate sample set comprising a union of at least a part of
8 said plurality of synthesized data and said plurality of real sample data, said
9 candidate sample set having a starting population, said candidate sample set
10 being unsupervised as to which members will be classified by said method as
11 being outliers;
 - 12 generating a set of classifiers, each member of said set being a procedure
13 or a representation for a function classifying an operand data as an outlier or a
14 non-outlier, said generating a set of classifiers including:
 - 15 initializing said set of classifiers to be an empty set,
 - 16 selectively sampling said candidate sample data to form a learning
17 data set, said selectively sampling including:
 - 18 i) applying said set of classifiers to each of said candidate
19 sample data and, if any classifiers are extant in said set, generating
20 a corresponding set of classification results,
 - 21 ii) identifying a consistency, for each of said candidate
22 sample data, among said data's corresponding set of classification
23 results.

iii) calculating an uncertainty value for each of said candidate sample data based on said identified consistency of said data's corresponding set of classification results.

iv) ii) calculating a sampling probability value for each of said candidate sample data based, at least in part, on the corresponding uncertainty value based on said set of classification results, and

v) iii) sampling from said candidate data to form said learning data set based, at least in part, on said sampling probability values, such that said learning data set has a population substantially lower than the starting population.

generating another classifier based on said learning data set, updating said set of classifiers to include said another classifier.

and

repeating said selectively sampling, said generating another classifier, and said updating until said set of classifiers includes at least t members; and

generating an outlier detection algorithm based, at least in part, on at least one of said another classifiers, for classifying a datum as being an outlier or a non-outlier.

24. (Currently Amended) The method of claim 23, further comprising:

receiving a given mis-classifying mis-characterizing cost data associated with at least one of said synthesized samples, representing a cost of said outlier detection algorithm mis-classifying mis-characterizing said at least one synthesized sample as a non-outlier.

25. (Currently Amended) The method of claim 24, wherein said constructing said learning data set from said candidate sample data is further based on said mis-classifying mis-characterizing cost.

1 26. (Currently Amended) The method of claim 23, wherein said calculating a
2 sampling probability value includes identifying a consistency within said set of
3 classification results and calculating a probability of said identified consistency
4 uses using a Binomial probability function.

1 27. (Currently Amended) The method of claim 23, wherein said calculating a
2 sampling probability value includes identifying a consistency within said set of
3 classification results and calculating a probability of said identified consistency
4 uses using a Gaussian probability function.

1 28. (Currently Amended) The method of claim 23, said machine-readable code
2 for generating synthesized data generates said synthesized data in accordance
3 with a given statistical likelihood of said generated data meeting an outlier
4 criterion.

1 29. (Previously Presented) The method of claim 23, wherein said generating an
2 outlier detection algorithm generates the outlier detection algorithm such that
3 said algorithm applies an aggregate of members of said set of classification
4 algorithms, calculates a corresponding set of detection result data representing
5 each of said aggregate's member's classification, and applies a voting scheme to
6 said corresponding set of detection result data.

1 30. (Currently Amended) The method of claim 23, wherein said calculating of
2 said uncertainty value includes identifying a consistency among said set of
3 classification results, and calculating a probability of said consistency assumes,
4 assuming each classification result within said set has a 50-50 probability of
5 representing an operand as meeting an outlier said alarm criterion, statistically
6 independent of said operand and of all other classification results within said set.

1 31. (Currently Amended) The method of claim 30, wherein said calculating-a
2 sampling probability value further includes calculating a probability of said
3 identified consistency uses using a Binomial probability function.

1 32. (Currently Amended) The method of claim 30, wherein said calculating-a
2 sampling probability value further includes calculating a probability of said
3 identified consistency uses using a Gaussian probability function.

1 33. (Currently Amended) The method of claim 28, wherein said calculating-of
2 said uncertainty value includes identifying a consistency among said
3 classification results, and calculating a probability of said consistency, assuming
4 assumes each classification result within said set has a 50-50 probability of
5 representing an operand as meeting said outlier alarm criteria, statistically
6 independent of said operand and of all other classification results within said set.

1 34. (Currently Amended) The method of claim 33, wherein said calculating-a
2 sampling probability value further includes calculating a probability of said
3 identified consistency uses using a Binomial probability function.

1 35. (Currently Amended) The method of claim 33, wherein said calculating-a
2 sampling probability value further includes calculating a probability of said
3 identified consistency uses using a Gaussian function.

1 36. (Currently Amended) A system for classifying externally detected samples as
2 one of at least normal and an outlier, comprising:
3 a machine controller having a readable storage medium;
4 a machine-readable program code, stored on the machine-readable
5 storage medium, having instructions to:
6 generate a plurality of synthesized data, each representing a randomly
7 generated state within a given vector space, said generating including generating
8 a random number generation;

9 receive a plurality of real sample data, each representing an observed
10 event as represented in said given vector space;

11 form a candidate sample set comprising a union of at least a part of said
12 plurality of synthesized data and said plurality of real sample data, said candidate
13 sample set having a starting population;

14 generate a set of classifiers, each member of said set being a procedure
15 or a representation for a function classifying an operand data as a rare event or
16 an outlier or a non-outlier, said generating a set of classifiers including:

17 initializing said set of classifiers to be an empty set,

18 selectively sampling said candidate sample data to form a learning
19 data set, said selectively sampling including:

20 i) applying said set of classifiers to each of said candidate
21 sample data and, if any classifiers are extant in said set, generating
22 a corresponding set of classification results, and

23 ii) identifying a consistency, for each of said candidate
24 sample data, among said data's corresponding set of classification
25 results,

26 iii) calculating an uncertainty value for each of said candidate
27 sample data based on said identified consistency of said data's
28 corresponding set of classification results,

29 iv) ii) calculating a sampling probability value for each of said
30 candidate sample data based, at least in part, on the corresponding
31 uncertainty value based on said set of classification results, and

32 v) iii) sampling from said candidate data to form said learning
33 data set based, at least in part, on said sampling probability values,
34 such that said learning data set has a population substantially lower
35 than the combined first and second population,
36 generating another classifier based on said learning data set,
37 updating said set of classifiers to include said another classifier,
38 and

1 46. (Currently Amended) The system of claim 41, wherein said machine-readable
2 code instructions for calculating of said uncertainty value include instructions for
3 identifying a consistency among said set of classification results, and calculating
4 a probability of said consistency, assuming that assume each classification result
5 within said set has a 50-50 probability of representing an operand as meeting
6 said outlier criterion alarm criteria, statistically independent of said operand and
7 of all other classification results within said set.

1 47. (Currently Amended) The system of claim 45, wherein said machine-readable
2 code instructions for calculating a sampling probability value include instructions
3 for calculating a probability of said identified consistency include instructions for
4 calculating said probability using a Binomial probability function.

1 48. (Currently Amended) The system of claim 45, wherein said machine-readable
2 code instructions for calculating a sampling probability value include instructions
3 for calculating a probability of said identified consistency include instructions for
4 calculating said probability using a Gaussian probability function.